

## Cloning of an Avilamycin Biosynthetic Gene Cluster from *Streptomyces viridochromogenes* Tü57

SIBYLLE GAISSER,<sup>1</sup> AXEL TREFZER,<sup>1</sup> SIGRID STOCKERT,<sup>1</sup> ANDREAS KIRSCHNING,<sup>2</sup>  
 AND ANDREAS BECHTHOLD<sup>1\*</sup>

*Pharmazeutisches Institut, Pharmazeutische Biologie, Universität Tübingen, Tübingen,<sup>1</sup> and Institut für Organische Chemie, Technische Universität Clausthal, Clausthal,<sup>2</sup> Germany*

Received 8 May 1997/Accepted 13 August 1997

A 65-kb region of DNA from *Streptomyces viridochromogenes* Tü57, containing genes encoding proteins involved in the biosynthesis of avilamycins, was isolated. The DNA sequence of a 6.4-kb fragment from this region revealed four open reading frames (ORF1 to ORF4), three of which are fully contained within the sequenced fragment. The deduced amino acid sequence of AvIM, encoded by ORF2, shows 37% identity to a 6-methylsalicylic acid synthase from *Penicillium patulum*. Cultures of *S. lividans* TK24 and *S. coelicolor* CH999 containing plasmids with ORF2 on a 5.5-kb *Pst*I fragment were able to produce orsellinic acid, an unreduced version of 6-methylsalicylic acid. The amino acid sequence encoded by ORF3 (AviD) is 62% identical to that of StrD, a dTDP-glucose synthase from *S. griseus*. The deduced amino acid sequence of AviE, encoded by ORF4, shows 55% identity to a dTDP-glucose dehydratase (StrE) from *S. griseus*. Gene insertional inactivation experiments of *aviE* abolished avilamycin production, indicating the involvement of *aviE* in the biosynthesis of avilamycins.

The avilamycins (Fig. 1), which are produced by *Streptomyces viridochromogenes* Tü57, are oligosaccharide antibiotics and belong to the orthosomycin group of antibiotics (10). Avilamycins as well as other important members of the orthosomycins contain a dichloroisoevernic acid moiety, as well as one or more orthoester linkages which are associated with carbohydrate residues (35). The compound SCH27899 shows excellent activity against gram-positive bacteria (22, 32) and is presently being tested for possible use against human infectious diseases (31, 37). Avilamycins inhibit the growth of gram-positive bacteria, and avilamycin A is a translation inhibitor binding to the 30S ribosomal subunit (34), but the exact mode of action of the avilamycins is not known. Avilamycins are used as an additive for animal breeding (MaxusG; Eli Lilly, Bad Homburg, Germany).

Few genetic studies have been carried out on the biosynthesis of orthosomycins. In 1992, Bergh and Uhlen (5) described the cloning and analysis of a polyketide synthase encoding gene cluster of *S. curacoi*, the producer of curamycin. The isolated gene cluster may be involved in the biosynthesis of curamycin or possibly in the biosynthesis of a spore pigment. Besides polyketide synthase genes, no other genes of this cluster were described. We recently reported a PCR method to amplify gene fragments coding for deoxynucleoside diphosphate (dNDP)-glucose 4,6-dehydratases (11), which are involved in the formation of 6-deoxyhexose moieties of different antibiotics (24). A PCR fragment was obtained by using chromosomal DNA from *S. viridochromogenes* Tü57 as the template. The deduced amino acid sequence of the fragment revealed similarity to known dNDP-glucose dehydratases (11). We have now used this PCR fragment as a probe to screen a cosmid library. On a cosmid hybridizing to the probe, three genes (*aviD*, *aviE*, and *aviM*) were detected. The disruption of

*aviE* affected avilamycin production. Expression of the multifunctional gene *aviM* in *S. lividans* TK24 or *S. coelicolor* CH999 resulted in the production of orsellinic acid. These data confirm that the cloned genes are part of the avilamycin biosynthetic gene cluster.

### MATERIALS AND METHODS

**Bacterial strains, plasmids, and materials.** *S. viridochromogenes* Tü57 and *S. lividans* TK24 were obtained from the culture collection of H. Zährner and W. Wohlleben, University of Tübingen (Tübingen, Germany). *S. coelicolor* CH999 was from D. A. Hopwood (Norwich, United Kingdom). Cosmid pOJ446 (7) was obtained from B. E. Schoner (Lilly Research Laboratories, Indianapolis, Ind.). Plasmid pWHM3 (33) was obtained from H. Decker (Hoechst, Frankfurt, Germany). Plasmid pBluescript-SK<sup>+</sup> (pSK<sup>+</sup>) was from Stratagene (Heidelberg, Germany). Medium components were purchased from Difco Laboratories (Detroit, Mich.), soya flour was purchased from Hartge Ingredients (Hamburg, Germany), and restriction enzymes were purchased from Amersham Life Science (Buckinghamshire, United Kingdom). Apramycin and avilamycins were a gift from Eli Lilly, carbenicillin was from Roth (Karlsruhe, Germany), erythromycin was from Fluka (Neu Ulm, Germany), and thiostrepton was from Sigma (Deisenhofen, Germany). All other chemicals were from Roth.

**Culture conditions.** *S. viridochromogenes* Tü57, *S. coelicolor* CH999, and *S. lividans* TK24 were maintained on HA medium containing 1% malt extract, 0.4% yeast extract, 1.6% agar, 0.4% glucose, and 1 mM CaCl<sub>2</sub> (pH 7.2) at 28°C (*S. viridochromogenes* Tü57 at 37°C). For the production of avilamycins, spores of *S. viridochromogenes* Tü57 were transferred to NL19+ medium containing 2% mannitol, 2% soya flour, and 20 mM L-valine (pH 7.2) and grown at 28°C in 500-ml baffled flasks filled with 100 ml of medium at 180 rpm. For the production of orsellinic acid, strains were grown on R5 agar plates for 5 days (4, 17). For the preparation of protoplasts, *S. viridochromogenes* Tü57 was cultivated in a modified S medium containing 0.4% peptone, 0.4% yeast extract, 0.4% K<sub>2</sub>HPO<sub>4</sub>, 0.2% KH<sub>2</sub>PO<sub>4</sub>, and 0.75% L-glycine (23) for 50 h. For protoplast preparation, *S. lividans* TK24 and *S. coelicolor* CH999 were grown in liquid R2YE for 30 h according to the standard procedure (17). All protoplasts were regenerated on R2YE medium. General methods for the cultivation of *Escherichia coli* XL1-Blue-MRF<sup>+</sup> were as previously described (27). Thiostrepton (0.025 mg/ml), carbenicillin (0.05 mg/ml), apramycin (0.05 mg/ml), and erythromycin (0.05 mg/ml) were used for selective growth of recombinant strains.

**DNA isolation, manipulation, and cloning.** Plasmid isolation, DNA endonuclease restriction analysis, ligation, and transformation were done by standard procedures (17, 27). Large-scale *E. coli* plasmid DNA was isolated with the Nucleobond AX100 kit (Macherey and Nagel, Düren, Germany). Genomic DNA of *S. viridochromogenes* Tü57 was isolated as described by Altenbuchner and Cullum (1). For the construction of a cosmid library from *S. viridochromogenes* Tü57, chromosomal DNA was partially digested with *Sau*3A, and fragments of 25 to 40 kb were ligated into the *Bam*HI site of cosmid pOJ446. DNA was packed

\* Corresponding author. Mailing address: Pharmazeutisches Institut, Pharmazeutische Biologie, Universität Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany. Ph ne: (0049-7071) 2975483. Fax: (0049-7071) 295250. E-mail: bechthold@uni-tuebingen.de.

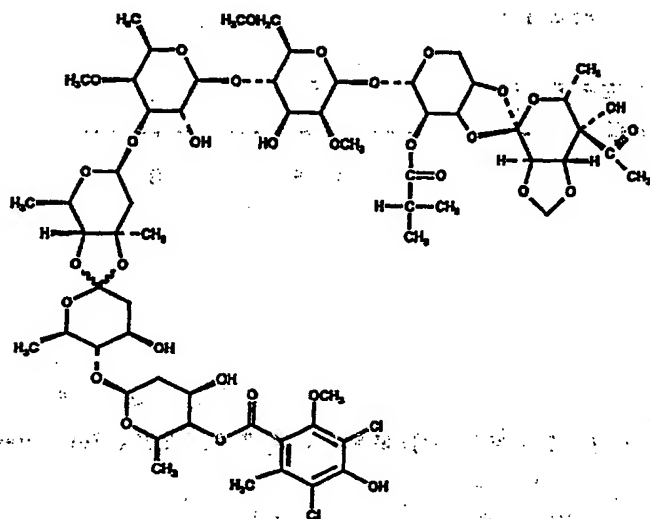


FIG. 1. Structure of avilamycin A.

into phages by using the Gigapack Packaging Extract Gold system from Stratagene. The phages were used to transduce *E. coli* XL1-Blue-MRF<sup>+</sup>. For screening of the cosmid library, a DNA fragment obtained by PCR amplification (11) was used as a probe. The probe was labeled with digoxigenin (DIG) by using a DIG labeling and detection kit (Boehringer, Mannheim, Germany). Colony and Southern hybridization were performed with Hybond N nylon membranes (Amersham) according to standard protocols (27).

**Sequencing.** DNA sequencing was performed on double-stranded templates (subclones in pSK<sup>+</sup>) by the dideoxynucleotide chain termination method, using a Thermo Sequenase core sequencing kit with 7-deaza-dGTP (RPN2440) from Molecular Dynamics (Krefeld, Germany). Both strands were sequenced with standard primers (T3, T7, M13 reverse, and M13 forward) or with internal oligonucleotide primers on a Molecular Dynamics Vistra 725 DNA sequencer or on an Applied Biosystems sequencer (model 377).

**Computer-assisted sequence analysis.** Computer-assisted sequence analysis was carried out by using the DNASIS software package (version 2, 1995; Hitachi Software Engineering, San Bruno, Calif.). BlastX analyses (2) were run with the GenBank CDC translations + PDB + SwissProt + SPupdate + PIR, release 3.17, 1997. Open reading frames were identified by using the CODONPREFERENCE program (12).

**Insertional inactivation of *aviE*.** To determine whether *aviE* is involved in avilamycin biosynthesis, insertional inactivation experiments were carried out. A 1.4-kb *EcoRI*-*BglII* fragment from *S. viridochromogenes* Tü57 containing the *aviE* gene was subcloned into pSK<sup>+</sup> which had been restricted with *EcoRI* and *Bam*HI. Into an internal *Sna*BI site of the 1.4-kb fragment, a 1.6-kb *Acc65I* fragment from pJ4026 (6, 30) carrying the *ermE* gene was subcloned to create pDesery<sup>+</sup>. Integration of pDesery<sup>+</sup> into the chromosome of *S. viridochromogenes* Tü57 was carried out by polyethylene glycol-induced protoplast transformation (17). Approximately 15 µg of single-stranded plasmid DNA, obtained by alkaline treatment, was used for transformation (16).

**Expression of the orsellinic acid synthase in *S. coelicolor* CH999 and *S. lividans* TK24.** A 5.5-kb *PstI* fragment containing the entire *aviM* gene was cloned into the *PstI* site of pWHM3. The fragment was inserted in both orientations to give either MSS4.3 (transcribed from the promoter of the thiostrepton resistance gene of pWHM3) or MSS4.5 (opposite orientation). These constructs were used to transform *S. coelicolor* CH999 and *S. lividans* TK24.

**Detection of avilamycins.** Cultures (400 ml) of *S. viridochromogenes* Tü57 and mutants obtained by insertional inactivation experiments were grown for 72 h in NL19+ medium and harvested by centrifugation. The medium was extracted with an equal volume of ethyl acetate, and cells were extracted with an equal volume of methanol. Ethyl acetate and methanol were removed under vacuum, and the products were combined and resuspended in a small volume of methanol. The crude product was chromatographed on a Sephadex LH20 column with methanol as the solvent. Fractions were analyzed by thin-layer chromatography (TLC) analysis on silica plates (Merck, Darmstadt, Germany) with CH<sub>2</sub>Cl<sub>2</sub>-methanol (9:1). Developed plates were sprayed with anisaldehyde solution (1% anisaldehyde in methanol-acetic acid-sulfuric acid [8:1:1]) and heated to 120°C for 5 min. The avilamycins turned black. Further analysis was carried out by high-performance liquid chromatography (HPLC) with a diode array detector, using a Nucleosil 100 C<sub>18</sub> column (5 µm) and a linear gradient (15 min) from 0 to 100% acetonitrile in 0.1% aqueous phosphoric acid (flow rate, 2 ml/min) (9, 14, 21). The minimal detectable concentration of avilamycin A was 0.01 mg/liter of medium. Extracts were also assayed for avilamycins by the agar diffusion technique, using *Bacillus subtilis* as the test organism (9).

**Detection of orsellinic acid.** Cultures (agar plates) of *S. coelicolor* CH999 and *S. lividans* TK24 containing pMSS4.3 were grown for 5 days on R5 medium and extracted with an equal volume of methanol. The solvent was removed under vacuum, and the product was resuspended in ethyl acetate. The organic layer was extracted with 1% aqueous Na<sub>2</sub>CO<sub>3</sub>. The aqueous phase was adjusted to pH 2 to 3 and was extracted again with ethyl acetate. The organic layer was concentrated in vacuum, and orsellinic acid was detected by TLC (see above). Orsellinic acid turned red after treatment with anisaldehyde solution. For the preparative isolation of orsellinic acid, 20 agar plates (approximately 400 ml) containing R5 medium were inoculated, cultivated for 5 days, and then extracted with 400 ml of methanol. The solvent was removed, and the product was resuspended in 200 ml of H<sub>2</sub>O-acetic acid (99:1). After extraction with 200 ml of ethyl acetate-acetic acid (99:1), the organic phase was extracted with 400 ml of 1% aqueous Na<sub>2</sub>CO<sub>3</sub>. The pH of this Na<sub>2</sub>CO<sub>3</sub> solution was adjusted to 2 to 3, and orsellinic acid was extracted with 400 ml of ethyl acetate. The solvent was removed, and products were resuspended in 10 ml of methanol. Further purification was achieved by preparative TLC on silica gel, using dichloromethane-methanol-acetic acid (9:1:0.1) followed by fractionation on a Sephadex LH20 column (40 by 2.5 cm), using methanol as the solvent.

**Nucleotide sequence accession number.** The nucleotide sequence data reported in this paper appear in the EMBL nucleotide sequence database under accession no. Y11985.

## RESULTS

**Screening of a cosmid library for the putative avilamycin biosynthetic gene cluster.** A cosmid library from *S. viridochromogenes* Tü57 DNA was prepared, and approximately 2,000 colonies were probed by colony hybridization using an internal fragment of a dNTP-glucose 4,6-dehydratase gene as a probe. This probe had been derived by PCR amplification using genomic DNA of *S. viridochromogenes* Tü57 as a template (11). Six colonies hybridized to the probe. Cosmid DNA isolated from these colonies was analyzed by restriction mapping and Southern hybridization experiments. All cosmids contained overlapping DNA encompassing in a total of approximately 65 kb of *S. viridochromogenes* Tü57 genomic DNA. One cosmid, F4, was used for further restriction mapping and sequencing (Fig. 2).

**Sequence analysis of a 6.4-kb region.** An dNDP-glucose 4,6-dehydratase gene is thought to be involved in the biosynthesis of carbohydrate moieties of avilamycins (24). It might therefore be expected that genes encoding biosynthetic enzymes for avilamycins would occur clustered with the dNDP-glucose 4,6-dehydratase gene. A 6.4-kb DNA fragment containing the dehydratase gene (fragment I [Fig. 2]) was sequenced. Within this segment, four open reading frames (ORFs) with the characteristics of *Streptomyces* genes (overall G+C content, 71.5%; high bias toward G and C in the third codon position) were identified. All four ORFs are transcribed in the same direction. One ORF is truncated, whereas the other three are fully contained within the sequenced fragment. The region between ORF2 and ORF3 is very rich in AT, and some parts can be viewed as a possible promoter sequence (8, 29) (Fig. 3).

**Deduced functions of the proteins.** The gene product of ORF1 (truncated) shows homology to DpsC from *S. peucetius* ATCC 29050 (15) and to ORFC from *Streptomyces* sp. strain C5 (36). The exact functions of these genes, which are involved in the biosynthesis of doxorubicin and daunomycin, respectively, are not known. However, it has been speculated that DpsC is involved in selecting propionyl coenzyme A (propionyl-CoA) as the starter unit for daunorubicin biosynthesis (26). The deduced amino acid sequence encoded by *aviM* (ORF2) exhibits similarity to type I polyketide synthases. The highest identity was found for a 6-methylsalicylic acid synthase (MSAS) from *Penicillium patulum* (3) (37% identity). A comparison between the amino acid sequences encoded by *aviD* and *strD* from *S. griseus* reveals 62% identity, suggesting that *aviD* encodes a dTDP-glucose synthase. The deduced amino

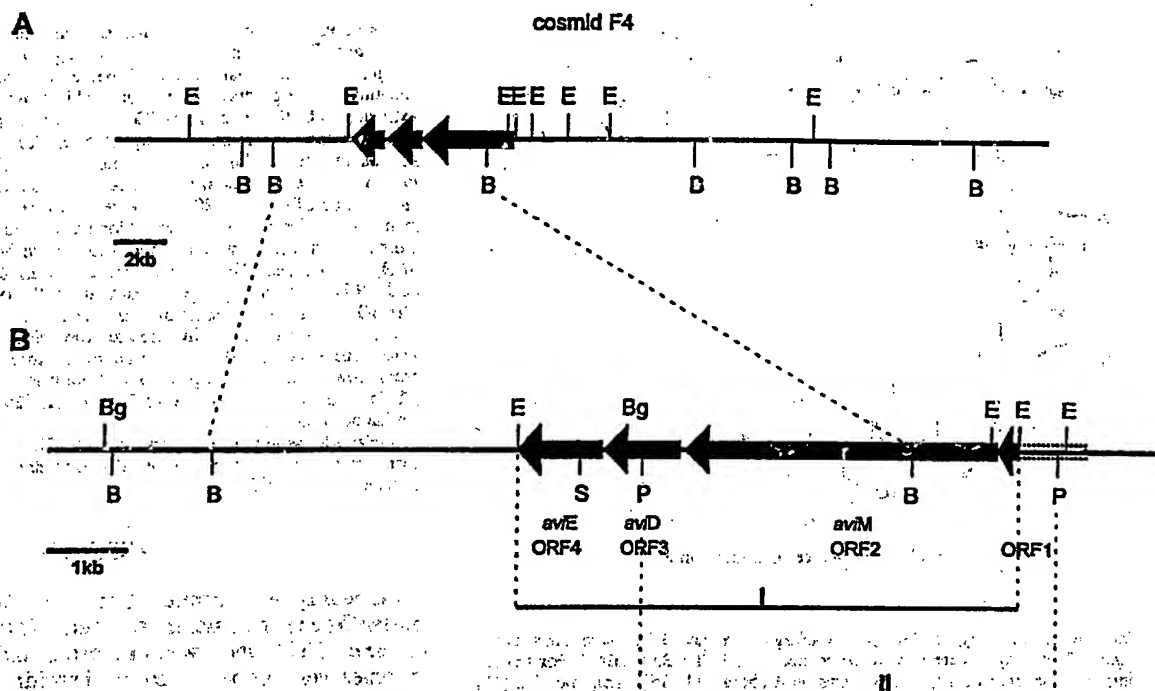


FIG. 2. (A) Restriction map of cosmid F4, containing DNA from *S. viridochromogenes* Tü57. (B) Detailed map of a 14-kb DNA fragment. ORF1, ORF2 (*avM*), ORF3 (*avD*), and ORF4 (*avE*) are marked as arrows. Line I indicates the sequenced region. The fragment which has been expressed in *S. coelicolor* CH999 is shown as line II. B, *Bam*HI; E, *Eco*RI; Bg, *Bgl*II; S, *Sna*BI; P, *Pst*I.

acid sequence of *aviE* shows 55% identity to the deduced amino acid sequence of *streE* from *S. griseus*, suggesting that *aviE* codes for a dTDP-glucose 4,6-dehydratase (13, 25).

**Insertional inactivation experiments.** To test whether we had in fact cloned genes responsible for avilamycin biosynthesis, we carried out insertional inactivation experiments using the vector pSK<sup>+</sup>, which does not replicate in *Streptomyces*. The 1.4-kb *Eco*RI-*Bgl*II fragment comprising *aviE* was subcloned into pSK<sup>+</sup>. Subsequently a 1.6-kb *Acc*65I fragment from pIJ4026 carrying the *ermE* gene was cloned into a *Sna*BI site located in the gene *aviE*, resulting in pDesery<sup>+</sup>. This plasmid was introduced into *S. viridochromogenes* Tü57 by protoplast transformation. Erythromycin-resistant colonies were obtained. These colonies were allowed to sporulate, and spores were selected again for resistance to erythromycin. Genomic DNA samples from different transformants were screened by hybridization, using a 1.6-kb *Acc*I fragment containing the gene *ermE* as a probe. Three mutants, in which genomic DNA hybridized to the probe, were selected. Genomic DNA derived from each transformant (-29, -7, and -1) was digested with either *Bam*HI or *Bgl*II and examined by Southern hybridization.

The 1.6-kb *ermE* fragment described above and the 0.5-kb PCR fragment containing parts of *aviE* were used as probes. The three mutants, but not the wild type, gave a hybridization signal with *ermE*, showing the integration of pDesery<sup>+</sup>. This result was confirmed by hybridization signals detected from all three mutants with a fragment of the vector pSK<sup>+</sup> (data not shown). Integration of pDesery<sup>+</sup> by a single crossover event can take place in two different ways, as depicted in Fig. 4B. Using *aviE* as a probe, crossover between the *Sna*BI and the *Bgl*II site will result in pattern A, which should give hybridization signals at 10.5 and 4.3 kb after *Bam*HI digestion. Crossover between the *Sna*BI and the *Eco*RI site results in pattern B, from which bands at 9.4 and 5.8 kb are expected. *Bgl*II

digestion should give identical 11-kb signals from all three mutants. As shown in Fig. 4A, the Southern plots were in accordance with these expectations, showing that the genotype of mutant -1 corresponded to pattern A, whereas those of mutants -29 and -7 corresponded to pattern B. None of the three mutants showed a pattern consistent with a double crossover.

**Analysis of the phenotypes of mutants -29, -7, and -1.** Mutants -29, -7, and -1 and the wild type were grown in liquid medium for 72 h. Extracts of strains were analyzed by TLC and HPLC as described in Materials and Methods. Different avilamycins, all identified by their characteristic UV spectra ( $\lambda_{max}$  at 214 and 288 nm), were produced by the wild-type strain (total avilamycin content, approximately 20 mg/liter of medium). No avilamycin (<0.01 mg/liter medium) was detected in extracts of mutants -29, -7, and -1.

**Expression of *avim* in *S. lividans* TK24 and *S. coelicolor* CH999.** A 5.5-kb *Pst*I fragment containing the entire *avim* gene (fragment II [Fig. 2]) was ligated into pWHM3 to create MSS4.3, in which *avim* should be transcribed from the promoter of the thiostrepton resistance gene of pWHM3 and MSS4.5, which contains *avim* in the opposite direction to the thiostrepton resistance promoter. These constructs were used to transform *S. coelicolor* CH999 and *S. lividans* TK24. Cultures of *S. coelicolor* CH999 and *S. lividans* TK24 containing plasmids pMSS4.3, pMSS4.5, and pWHM3 were cultivated, and extracts were analyzed by TLC. A UV-fluorescent compound turning red after treatment with anisaldehyde solution was detected in transformants containing pMSS4.3 but not in the transformants containing pMSS4.5 or pWHM3. The compound was isolated from agar plates of *S. coelicolor* CH999 in preparative scale as described in Materials and Methods and was fully characterized by nuclear magnetic resonance (NMR) spectroscopy in *d*<sub>6</sub>-acetone and mass spectrometry. This aromatic compound was identified as orsellinic acid. The 400-

GAATTCGGGGGGGAGATGGCCACTGCGGGGCGAGGACCACTGCTGCTCATGGAACACATCGTACGGACGGGTGAGCTCGGGGGGGGGGACC 95  
 NSAGRSATCGASDHLSMESHIVRTGELAPGDH  
 ORF1 (end)  
 ACGTCTGCACTGCCACGGCCCGGGGCTGGTGTGTCCAGGGCGCTCTTCAAGTTCTCGAATCCCGGACTGGGACGATGATGAATACCG 190  
 VLQLATAPQLVVS SA VLQVLESF DNDSE  
 ORF2 (Xv18) →  
 GCAAGCAGAACCGAATATGCCCGTGATATCCCTTGATCGCCCGGGGACTCCCTCTCTCGGTGGCCGTGGTGGGATCGGATGCCGGTTCGCC 285  
 NDERNNMPVVISLDRPADSLSPYAVVGIQCRFP  
 GGTGGTGCAACTCGCCCGGGAATCTGGGCACTGTGTGACGGCGGAGCGCAACACCGTGGGAGATGCCCGGACCGCTGGGAGAGTACCG 380  
 GGVNSPCEPMDLLTAGRNTVGSMPDRWEEYR  
 GGACTTCGGTCCGGTTCGATCGCGGCTGGGAGCGGATTCGCTCGGCGASTTCTCTGGACGACACATAGCGGGTTTCGAGCGCGAGTTCT 475  
 DFGPFPDAALSTAIRS GSFLLDDIAGF DAEFF  
 TGGGATCTCGCCCGGAGCGGAGCTGATGACCGCGACAGCGCTGATGCTGGAGTGGCTGGCAGGCACTGAGGACCGCGGATCCCG 570  
 GISPREAELMDPQORLMLLEVANQALEHAGIF  
 CCGCAGCCCTGGCAGGACCGACCGGTGTGTTCGCGGGTCTGACCTAGACTATGGCGCGGAGCGGTGGAGAGCTCGGACACATCGA 665  
 PHTLAGTDTGTGVPAGVCTVDYTGAGRLDLPNID  
 CCGATGGACCGGATCGCGCGCGGCTGCGCGCGTGTCCAAACCGCTCTCTACCGCTCGACCTCGCGCGCGCGCTCTCTCATGACACCG 760  
 AWTGIGA AVCAVSNVSHALDLRQPSLSIDTA  
 CGTCCGCGCGCTCGCTGGTGGCTGACACCGCGCGAGCGCTCGCGTGGCGAGTGCACCTGGCGCTGGCGCGCGCTCGCGCGCTCGCTG 855  
 CSASLVALHTAAQSRLRGECTVALAGGVHLL  
 GTGTCGCGCGCGAGCATCGCTGGCAGCGCGCGCGCTGGCGCGCGCGAGCAGCGCTGGCGCGCTGGCGCGCGCGCGCGCGCTAGCC 950  
 VSPGQTIALGTAGALAPDQGRSKPFAAASG  
 CGTTCGCGCTGGCGCGCTAGCGCGCGCGCGCGCGCTGGCGCGCTGGCTCATAGCTGCTAGCGCGCGCGCTAGCGCGCGCGCGCGCGCT 1045  
 VAASADGYGRGEGCGVLLVIXLLTDVARDGRV  
 TGTGCTGCTGGCGCGCGCGCTTCAACCGAGCG 1140  
 LAVLRGS AFM QDORTNGIMAPCGQQAQEHVR  
 CGCGCGCTGGCGCGCGCGCTGGCTGGCGCGCGCGCTGGCTGGCG 1235  
 RALTAAGVAADTVDFV EANGTGT RLGDPMHIG  
 CGCGCTGGCGCGCGCTGGCG 1330  
 AIAAVYGRDRSGQEP CAVGVSGVSHGLGSA  
 CGCGCGCTGGCGCGCGCTGGCG 1425  
 G VAGV IKA ILALDEDRIPASLLD DDFHFEID  
 TGGCGCGCGCTGGCG 1520  
 MAGLDIRLATRALPMPERFERRAAVSGFGYG  
 CGCGCGCTGGCGCGCGCTGGCG 1615  
 GTVAHVVL EAGTAPARPAFFAGGLFPVSA  
 CCGCGCGCTGGCGCGCGCTGGCG 1710  
 SPEALDRRAALAEERVEEGADLASVGH TLAH  
 CGCGCGCTGGCGCGCGCTGGCG 1805  
 RQSPFLVHRAAVVATGDELAAGLEALATQEP  
 CGCGCGCTGGCGCGCGCTGGCG 1900  
 FGLVTGALLPDAGCRFVWVFSGHGSGMAGMGR  
 AACTCTGGAGCGCGCGCTGGCG 1995  
 LLEAEFPVFAEVIDELSPVPKESBIGFSPRQHL  
 CTGGAGGCGCGCGCTGGCG 2090  
 LEGDHTHEVDGAQTNTFAMQLGLAALWRSRGV  
 CGCGCGCTGGCGCGCGCTGGCG 2185  
 PAAVIGHSVGEETAAAVTAGALTVTGDARLICR  
 GTAGCTGCTGCTGGCG 2280  
 RSZLLRRAAGRGANAMVSLPFDAAERLAGN  
 GACCGCGCTGGCGCGCGCTGGCG 2375  
 DAVVAAIAASCTTSTTSTVISGDPGEVEKVVGG  
 CGCGCGCTGGCGCGCGCTGGCG 2470  
 EGLVVRVAVSDVAFNSPHMDPLDLRLRAADH  
 AGCTGCTGGCGCGCGCTGGCG 2565  
 LSPSAPFHTPLYTALADPRAVTVTADGAYWAA  
 AACTCGCGCGCGCTGGCG 2660  
 HLRNFVRLAAALTAALAEDEGHRAPFVLSFPFVV  
 CGCGCGCTGGCGCGCGCTGGCG 2755  
 AHSIHETLAEERGVEDVTVGPTLRLRNQPEARTF  
 CGCGCGCTGGCGCGCGCTGGCG 2850  
 EAAVGAANECHGVSDVMSALQFNGHLLFPY  
 CCGTGGCAGCGCGCGCTGGCG 2945  
 PMQHEPLWRSLAGARAAREGH D VDSHTLGT  
 CGCGCGCTGGCGCGCGCTGGCG 3040  
 GGVAGSDLR LWNHSTLDDDSRFPYPGSHALGVR  
 AGATCTGGCGCGCGCTGGCG 3135  
 IVPAAVLAVTFLAAGAEGERRALQDMTMT  
 CGCGCTGGCGCGCGCTGGCG 3230  
 FVLTAGQRQIQVVRREGVRLASRTVADAADP  
 GAAACCGCGCTGGCGCGCGCTGGCG 3325  
 NPAWLVHAEARTAAPDLAAGLAARSLLDPGEHR  
 GGTGCGCGCGCGCGCGCGCTGGCG 3420  
 LEPADPGLVSRRLAEVGVSTGFDMSVRL  
 CGCGCTGGCGCGCGCTGGCG 3515  
 AGGLGVLR AOVLS PDAASMAFLLDVAVMSIAFAA  
 CTGCTGGCGCGCGCTGGCG 3610  
 FVGLPQLRMVHVHDEITVDGTTFEAAATVEVAL  
 TCGATCCCGCGCGCTGGCG 3705  
 DFRVADTVHALVTDGEGRPVAVSLRGLRYPVV  
 GAGCAGCG 3800  
 EQPAAPD TDEFGGDADADVVSFAQLSPFELRS  
 CGCGCTGGCGCGCGCTGGCG 3895  
 RVLDEVRERQIAQEMRLAAPTALHVRREPLVEQGL  
 TCGATCGGTGATACCTGACCGAGCGCTACGGAACCG 3990  
 DSVMTTVVVRERLEKELGRDVPANIFWKLPTI  
 AGCGCATCTGTGATACCTGACCGAGCGCTACGGAACCG 4085  
 SDIVDHLTERLTEREPTADGHA  
 CG 4180

FIG. 3. Nucleotide sequence for the analyzed region. ORF1, truncated to 179 (TGA); ORF2, 178 (ATG) to 4057 (TGA) (1,293 amino acids); ORF3, 4249 (ATG) to 5314 (TGA) (355 amino acids); ORF4, 5367 (ATG) to 6432 (TGA) (355 amino acids). ATG start codons and TGA stop codons are in boldface. Possible ribosome binding sites are underlined.



ORF3 (*aviD*) →

CATTGATGCCATCTGGAGTTCCTTGAACAGCACTGATCTGACGGCGCTGACCGCAAGGTGTGGGATATGAAAGGCGCTCATATTGGTGGCGGG 4275

M K A L I L V G G

GTTCCTGAGGCTGGCGGCGATTACGCATACCTTCGGCCAAACAACTGGTTCGGGTCGCAATAGAACTTCTCTCTCTATGATGATTCAT 4370

V G S R L R P L T H T S A K Q L V P V A N S T V L S Y V L D S I

CAAGGAGGCGGGAATATCGAGGTGGGATTTGCTGCTGGGGGACCGCGCGGAGATCCAGGCTTCCTGGGAGACGCGCGCGAGTTCGGGCTGG 4465

K E A G I S E V G I V V G A T A A E I Q A S V G D G A E F G L D

ACGTCACTTACATCCAGCAGGACCGCGCGCGCTGGCGGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4560

V T Y I Q Q D A P R G L A D A V L V S R D F L G D D D F V M Y

CTGGGCGACAACTAGCTGCTGGAGCGCATCTGCT 4655

L G D N Y V V D G I V D P R R D K P A A Q V M L A R V A

CGACCGCGCAGGTTTCGGGCGGAACTTGACGAGAACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4750

D P R R F G V A E L D E N G R V T A V V S K P Q D P R S D L A I

TCCTGCGCT 4845

V G V Y A S S A V H S A V A G I R P S M R N E L E T D A L

CAGTGGCTTCCTGAGCAGGCGCGGCT 4940

Q M L A D H G H E V G S T V I D R E Y N K D V A S V T D V L E M N

CGCGCACCTGCT 5035

R H L E S L E T A L D O R V D E A S E L V G R V V V A G A K

AGATCTCGCGCT 5130

I S G S R I V G P A V I G R D T V V Q D S Y L G P F T S I E G

ORF4 (*aviE*) →

CCCTCGGAGCGCGCGGACCGGACCGGCTGACCGGACCGGCG 5415

M R D S A G A R R G R H V L V I

TCCTGAGGCGCGCGCTCATCGCTCCCACTGCTGCGGAACTGCTGCGGACCGGCG 5510

G C A G F I G S Q Y V R E L V R D G D P A R V T V L D K L T Y

CGCGGAACTGCGGAACTGCGGAACTGCGGAACTGCGGAACTGCGGAACTGCGGAACTGCGGAACTGCGGAACTGCGGAACTGCGGAACTGCGGAACT 5605

A G M L A N L E F V A G R Y T F V H S D I C D A R L L A E V V F

CGCGCACTGCT 5700

G H D L V V N V A A E S H V D R S I A D A A P F I R T N N V Q G V

TCGAGGCTCATCGAGTTCCTGCGGCG 5795

Q A L M Q V C L E A G T A R I V Q V S T D E V Y G S I E T G S

TCGAGGCTCATCGAGTTCCTGCGGCG 5890

W D E D A P I A P N E Y A A S E A O G D M V A L A I A T H G

GCT 5985

L P V S V T R C G N H Y G P Y Q P F E K V V P L F T T R L D G

GCCACGCACTGCT 6080

H S I P L Y G D G G N Y R D W H V S D H V R G I R L V A E R

GCGCT 6175

G L P G Q V Y K I A G S A E L T H L E L T Q Q L D A L G A G M

GCGACGCT 6270

D R V E R V D R E G H D R R Y S L C D A R L G A L G Y A P R V

TGCT 6365

S E H Q G L A D T V R N Y A E N R D M W E P L N E R A L S D G

ATGCT

H A L A V P T A P K T R G E F A R G H F L K E

FIG. 3—Continued.

MHz  $^1\text{H}$  NMR spectrum shows two resonances for 5-H at  $\delta = 6.27$  ppm ( $d, J = 2.3$  Hz) and for 3-H at 6.21 ppm ( $d, J = 2.3$  Hz). The methyl group is located at 2.53 ppm and shows a large relative nuclear Overhauser effect (16%) with 5-H. In addition, three exchangeable protons at 6.20 to 5.53 ppm (b, 2H) and 3.32 ppm (s, 1H) were detected. The constitution of the compound was further proven by recording of a  $^{13}\text{C}$  NMR spectrum ( $\delta = 174.9$  [CO<sub>2</sub>H], 167.1 [C-4], 163.2 [C-2], 145.1 [C-1], 112.0 [C-5], 101.6 [C-3], 105.9 [C-6], 24.5 [CH<sub>3</sub>] ppm) and a MS(EI) spectrum ( $m/z = 168.0422$  [M<sup>+</sup>]).

Approximately 30 mg of orsellinic acid was produced from 20 agar plates (400 ml of medium), an amount comparable to the yield for 6-methylsalicylic acid production in *S. coelicolor* CH999 (4).

## DISCUSSION

Considerable progress has been made recently in clarifying the molecular genetics of polyketide antibiotic biosynthesis in actinomycetes. These studies have led to the development of novel polyketides, which were obtained by mixing biosynthetic genes of different biosynthetic clusters and functional expression of these genes in a special host strain (19, 20). One of the areas that remain to be explored is the molecular basis for the biosynthesis of glycosylated compounds and especially of oligosaccharide antibiotics. In our study, three genes located on a cosmid prepared from genomic DNA of *S. viridochromogenes* Tü57 have been isolated and sequenced. It has been shown previously that dNDP-hexose synthases and dNDP-hex-

ose 4,6-dehydratases are involved in the biosynthesis of several 6-deoxyhexoses (24). The strong resemblance of *aviD* and *aviE* to *strD* and *strE*, both involved in the biosynthesis of dTDP-4-keto-6-deoxyglucose in *S. griseus*, indicates that these genes are involved in early steps of the biosynthesis of 6-deoxyhexoses, which are components of the avilamycins. Insertional inactivation experiments confirmed the involvement of these genes in the biosynthesis of avilamycins. The integration of pDesery into the genome disrupted a transcription unit with *aviD*, *aviE*, and further genes located downstream to *aviE*. This resulted in the abolition of avilamycin production.

Polyketide synthases have traditionally been classified as iterative polyketide synthases and modular polyketide synthases. Iterative polyketide synthases have been classified as type I or type II on the basis of the structural organization of the enzymes. Type I enzymes consist of multifunctional proteins in which individual active sites occur as domains, as exemplified by an MSAS from the fungus *P. patulum*. Type II enzymes consist of several separate monofunctional proteins. These proteins are involved in the production of aromatic compounds such as actinorhodin, granaticin, and tetracenomycin. In contrast, modular polyketide synthases consist of several domains, each containing a set of active sites, required for one step in carbon chain assembly and modification. Modular polyketide synthases are exemplified by a 6-deoxyerythronolide B synthase from *Saccharopolyspora erythraea* (18).

The orsellinic acid synthase reported here, which is the gene product of *aviM*, is a multifunctional protein belonging to iterative polyketide synthase type I. The overall homology

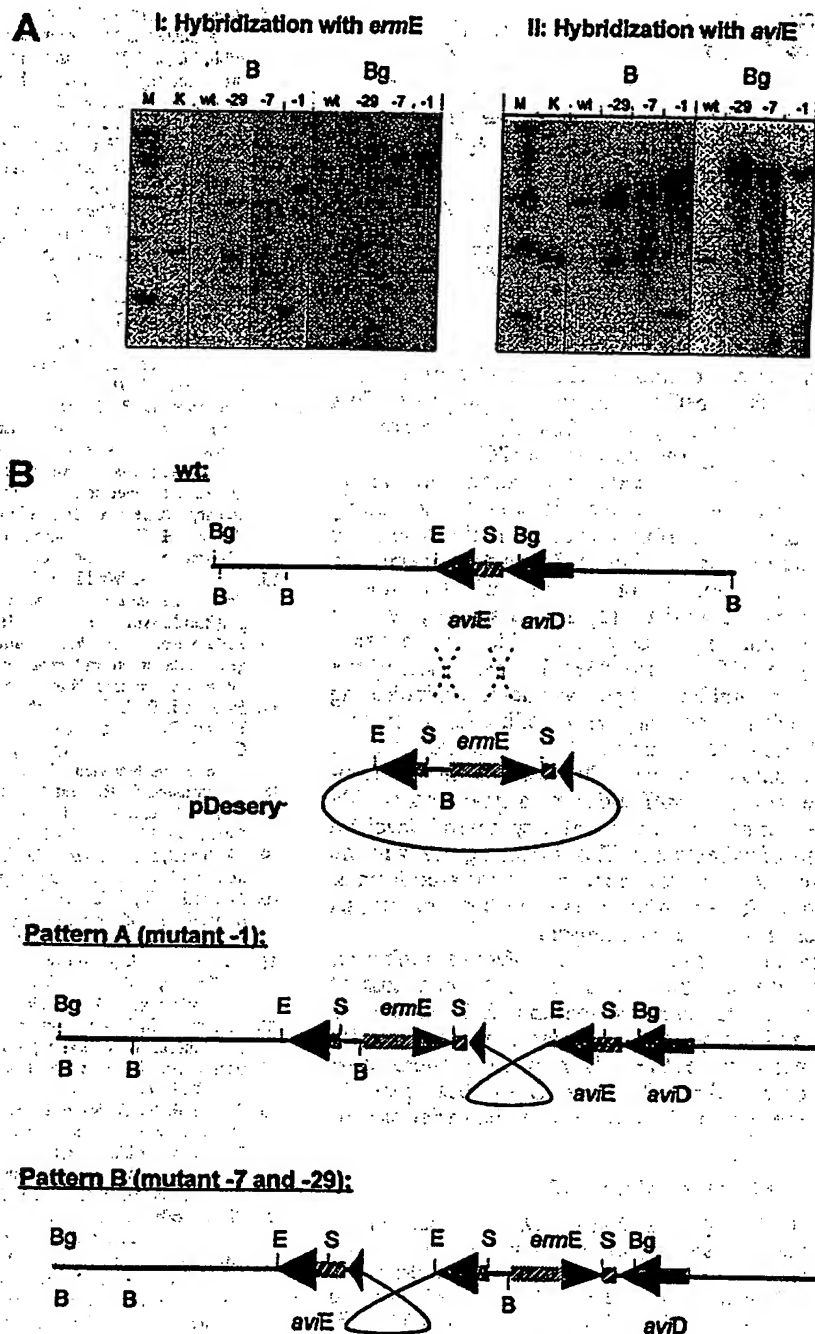


FIG. 4. (A) Southern hybridization of genomic DNA. Genomic DNA of the wild-type (wt) *S. viridochromogenes* Tu57 and of mutants -29, -7, and -1 was digested with *Bam*HI (B) and *Bgl*II (Bg) and probed with the *ermE* gene (I) and *aviE* (II). Lane M,  $\lambda$ -HindIII markers (DIG labeled); lane K, pDesery<sup>r</sup> restricted with *Bam*HI. (B) Possible patterns of integration of pDesery<sup>r</sup> into the genome of *S. viridochromogenes* Tu57 by single crossover events.

(37% identity) to the MSAS from *P. patulum* (3) is not very high but still significant. Motifs resembling acyl carrier proteins (ACPs),  $\beta$ -ketoacyl:ACP synthases, and acetyl-CoA/malonyl-CoA:ACP acyltransferase are detected in the orsellinic acid synthase; all of these motifs have also been detected in the MSAS (Fig. 5). As  $\alpha$ -keto reduction is necessary for the production of orsellinic acid, the absence of a ketoacyl reductase motif in the orsellinic acid synthase was expected. A further motif in the MSAS resembling dehydratases has not been described, but it might be located between amino acids 1216

and 1383 of this protein. This part shows some homology to the dehydratase motif in the 6-deoxyerythronolide B synthase 2 from *Saccharopolyspora erythraea* (data not shown), and as expected, this part is missing in *AviM* (Fig. 5). Unlike animal fatty acid synthases, MSAS and *AviM* do not harbor thioesterase domains.

The production of orsellinic acid after expression of *aviM* in *S. lividans* TK24 or *S. coelicolor* CH999 clearly confirmed the function of *aviM*. To our knowledge, this is the first type I polyketide synthase isolated from bacteria which is able to